The structure and biosynthesis of heinamides A1-A3 and B1-B5, antifungal members of the laxaphycin lipopeptide family

Heinilä, Lassi Matti Petteri

2021-07-07

Heinilä , L M P , Fewer , D P , Jokela , J K , Wahlsten , M , Ouyang , X , Permi , P , Jortikka , A & Sivonen , K 2021 , ' The structure and biosynthesis of heinamides A1-A3 and B1-B5, antifungal members of the laxaphycin lipopeptide family ' , Organic & Biomolecular Chemistry , vol. 19 , no. 25 , pp. 5577-5588 . https://doi.org/10.1039/d1ob00772f

http://hdl.handle.net/10138/333168
https://doi.org/10.1039/d1ob00772f

Organic & Biomolecular Chemistry

ROYAL SOCIETY OF CHEMISTRY

PAPER

View Article Online
View Journal | View Issue

Cite this: Org. Biomol. Chem., 2021, 19, 5577

The structure and biosynthesis of heinamides A1–A3 and B1–B5, antifungal members of the laxaphycin lipopeptide family†

Lassi Matti Petteri Heinilä, [a] David Peter Fewer, [a] Jouni Kalevi Jokela, [a] Matti Wahlsten, [a] Xiaodan Ouyang, [a] Perttu Permi,[b,c] Anna Jortikka [a] and Kaarina Sivonen *[a]

Laxaphycins are a family of cyclic lipopeptides with synergistic antifungal and antiproliferative activities. They are produced by multiple cyanobacterial genera and comprise two sets of structurally unrelated 11- and 12-residue macrocyclic lipopeptides. Here, we report the discovery of new antifungal laxaphycins from Nostoc sp. UHCC 0702, which we name heinamides, through antimicrobial bioactivity screening. We characterized the chemical structures of eight heinamide structural variants A1–A3 and B1–B5. These variants contain the rare non-proteinogenic amino acids 3-hydroxy-4-methylproline, 4-hydroxyproline, 3-hydroxy-D-leucine, dehydrobutyrine, 5-hydroxyl -amino octanoic acid, and O-carbamoyl-homoserine. We obtained an 8.6-Mb complete genome sequence from Nostoc sp. UHCC 0702 and identi ed the 93 kb heinamide biosynthetic gene cluster. The structurally distinct heinamides A1–A3 and B1–B5 variants are synthesized using an unusual branching biosynthetic pathway. The heinamide biosynthetic pathway also encodes several enzymes that supply non-proteinogenic amino acids to the heinamide synthetase. Through heterologous expression, we showed that (2S,4R)-4-hydroxy-L-proline is supplied through the action of a novel enzyme LxaN, which hydroxylates L-proline. 11- and 12-residue heinamides have the characteristic synergistic activity of laxaphycins against Aspergillus avus FBCC 2467. Structural and genetic information of heinamides may prove useful in future discovery of natural products and drug development.

Received 21st April 2021,
Accepted 29th May 2021

DOI: 10.1039/d1ob00772f

rsc.li/obc

Introduction

Cyanobacteria produce a wide range of bioactive natural products with unusual structures and potent bioactivity.[1,2] Knowledge of natural product chemical structures and biosynthetic mechanisms can facilitate their use in the pharmaceutical industry as new active compounds[3–5] and provide insights into their ecological role.[6] Laxaphycins are cyanobacterial natural products with two distinct macrocycles, 11- and 12-residue types, which act in synergy to produce antiproliferative and antifungal activities.[7,8] Laxaphycins include characteristic non-proteinogenic amino acids, including 3-hydroxy-D-leucine (OHLeu), 3-hydroxy-D-asparagine (OHAsn), dehydrobutyrine (Dhb), and (2S,4R)-4-hydroxyproline ((2S,4R)-4-OHPro).[9,10] Altogether 17 variants of 11-residue laxaphycins and 23 variants of 12-residue laxaphycins have been described from a broad range of cyanobacteria (Table S1†). Although 11- and 12-residue types are both found in most producer strains, there are reports of strains for which only one type of the compound is reported.[11,12]

In an earlier study, we described the biosynthetic pathway of scytocyclamides, which belong to the laxaphycin peptide family.[10] This pathway includes a shared initiating fatty-acyl AMP ligase (FAAL) and a polyketide synthase (PKS) module that branches with two non-ribosomal peptide synthetase (NRPS) pathways to produce the two distinct 11- and 12-residue compounds.[10] Scytocyclamides are produced by Scytonema hofmannii PCC 7110.

The aim of this study was to identify and describe new antifungal compounds from cyanobacteria. Members of the genera Candida and Aspergillus can cause invasive infections in humans, typically in immunocompromised patients.[13,14] Only a few chemical families of antimicrobials are currently used to treat fungal infections, and fungal resistance to these

[a]Department of Microbiology, Faculty of Agriculture and Forestry, University of Helsinki, Helsinki, Finland. E-mail: kaarina.sivonen@helsinki.fi
[b]Department of Chemistry, University of Jyväskylä, Jyväskylä, Finland
[c]Department of Biological and Environmental Science, Nanoscience Center, University of Jyväskylä, Jyväskylä, Finland

†Electronic supplementary information (ESI) available. See DOI: 10.1039/d1ob00772f

compounds is growing.[13–15] Cyanobacteria are known producers of antifungal compounds such as laxaphycins, hassallidins, nostofungicidine, and cryptophycins, which could be used as antifungal drug leads.[7,16–19]

Here, we identified novel members of the laxaphycin family of natural products, heinamides, through bioactivity-guided fractionation of Nostoc sp. UHCC 0702 extracts. Heinamides have antifungal activity that inhibit the growth of Aspergillus flavus FBCC 2467 with synergistic e ect between 11- and 12-residue type heinamides. We describe the chemical structures of heinamides A1–A3 and B1–B5 and identified the heinamide biosynthetic pathway. While the biosynthetic pathway is generally similar to the previously described scytocyclamide pathway, the di erences provide a broader view of the laxaphycin biosynthesis pathways. The heinamide biosynthetic pathway encodes enzymes for the production of the unusual amino acids (2S,4R)-4-OHPro and 3-hydroxy-4-methylproline (OHMePro), which appear in heinamide structures. The action of the proline hydroxylase LxaN from Nostoc sp. UHCC 0702 was shown through heterologous expression. A homolog of LxaN was found also in the genome of S. hofmannii PCC 7110, a producer of scytocyclamides, which also contain (2S,4R)-4-OHPro.

Experimental

Strains

Cyanobacterial strains used in this study were Nostoc sp. UHCC 0702, isolated from the Finnish freshwater lake Villähteen Kukkanen (60°57 19.4 N 25°53 02.4 E) in 19.8.2013, and S. hofmannii PCC 7110, which originated from a limestone cave in Bermuda.

Antimicrobial screening

Nostoc sp. UHCC 0702 was grown in 250 mL of Z8 medium at 20–21 °C with photon irradiation of 15 µmol m^{-2} s^{-1}. Erlenmeyer flasks of 500 mL were used with 250 mL of medium, with constant sterilized air bubbling for 3–5 weeks. The biomass was collected by decanting excess media and centrifugation at 8000g for 5 min. The biomass was frozen at 80 °C and freeze-dried with a CHRIST BETA 2–8 LSC plus freeze drier with a LYO CUBE 4–8 chamber. A total of 100 mg of freeze-dried Nostoc sp. UHCC 0702 biomass was extracted with 1 mL of methanol and glass beads (0.5 mm glass beads, Scientific Industries Inc, USA) using a FastPrep cell disrupter at 6.5 m s^{-1} two times for 25 s with a resting time of 5 min between runs. The samples were centrifuged at room temperature at 8000g for 5 min. The supernatant was collected and extraction of the biomass was repeated with 1 mL of methanol. The 2 mL combined supernatant was stored at 20 °C.

Antimicrobial activity screening was performed using 17 strains of fungi and bacteria (Table S2†). A total of 50 µL of cell extract, 50 µl of methanol (negative control) and 10 µL nystatin (5 mg mL^{-1}, Nystatin, Streptomyces noursei, EMD Millipore Corp, Germany) or 10 µL ampicillin (50 mg mL^{-1} in 70% ethanol, Ampicillin sodium salt, Sigma, Israel) were placed directly on the agar surface prior to inoculation with an indicator strain. Nystatin was used as a positive control for fungal assays while ampicillin was used as a positive control for bacterial assays. Solvents of the extract and controls were allowed to evaporate, leaving the solids di used in the agar. Inoculant was prepared by growing fungi for 2–14 days on potato dextrose agar (PDA) media at 28 °C and bacteria for 2 days on brain heart infusion (BHI) agar at 37 °C. Inoculant cell mass was transferred with a cotton swab from the agar to 3 mL of sterile 5 M NaCl solution, or sterile water in the case of A. flavus. Solution was spread on the assay plate with a fresh cotton swab. Fungal plates were incubated at 28 °C and bacterial plates at 37 °C for 2 days and examined for the presence of inhibition zones.

Disc-di usion assays were performed with purified heinamide for more quantitative analysis of bioactivity. Paper discs (Blank monodiscs, Abtek biologicals Ltd, UK) were prepared with methanol solutions of the peptides, methanol as negative control, and nystatin as positive control. A. flavus inoculum was prepared as previously and spread on the plate. Disks were placed on agar, the plates were incubated at 28 °C for 2 days, and examined for the presence of inhibition zones.

Purification of heinamides

Mass cultures were grown in modified Z8 medium without added nitrogen at 20–21 °C with photon irradiation of 15 µmol m^{-2} s^{-1}. Five-liter Erlenmeyer flasks were used with 2.7 L of medium, with constant sterilized air bubbling for 3–5 weeks. Cells were collected by decanting excess media and centrifugation at 8000g for 5 min. The cells were frozen at 80 °C and freeze-dried as described earlier.

A total of 30 mL of methanol was used per 1 g of dry cells. Cells were homogenized with a Heidolph Silentcrusher M at 20 000 rpm for 30 s. The solution was centrifuged 10 000g for 5 min and supernatant was collected. The extraction was repeated with 30 mL of methanol using the cell pellet. Chomatorex (Fuji-Davison Chemical Ltd, Aichi, Japan) chromatography silica ODS powder (10 mL) was added to the supernatant pool and the mixture was dried with a rotary evaporator Büchi Rotavapor R-200 at 30 °C. Solid phase extraction (SPE) was performed with Phenomenex SPE strata SI-1 silica 5 g per 20 mL column, preconditioned with 20 mL isopropanol and 20 mL of heptane. Silica ODS powder with the dry extract was added on top of the column and extracted with heptane, ethyl acetate, acetone, acetonitrile, and methanol, 40 mL each, with every fraction collected individually. Fractions were dried with a nitrogen gas flow and re-dissolved in 1 mL of methanol for bioactivity assays.

The active methanol fraction was further fractionated with an Agilent 1100 Series liquid chromatograph with Phenomenex Luna C18(2) (150 × 10 mm, 100 Å) column. Sample was injected in 100 µL batches and eluted with acetonitrile/isopropanol 1 : 1 (solvent B) and 0.1% HCOOH (solvent A) with initial isocratic stage of 40% solvent B in A for 15 min, followed by a linear gradient of solvent B from 40% to 100% in 10 min with a flow rate of 3 mL min^{-1}. Four heinamide frac-

tions were collected, dried with nitrogen flow, and weighed. Fraction 1 contained heinamides B1 and B2 (1 : 1), fraction 2 contained heinamide B1, fraction 3 contained heinamides A1, B3, and B4 (7 : 2 : 1) and fraction 4 contained heinamide A2. To further separate the products, fraction 3 was treated with an additional HPLC run with isocratic conditions of 41% solvent B in solvent A for 30 min with flow rate of 3 mL min $^{-1}$ using the same column. Fraction 3 was thus separated to fractions 3a and 3b containing heinamides B3 + B4 and A1, respectively. Heinamides A3 and B5 were not purified due to low production levels.

Amino acid analysis

To elucidate the stereochemistry of the eight heinamide structural variants, amino acid analysis was performed with the Marfey method as described before[20] with the isolated heinamides A1–A2 and B1–B4. The reagents 1-fluoro-2,4-dinitrophenyl-5-L-alanine amide (FDAA), 1-fluoro-2,4-dinitrophenyl-5-L-leucine amide (L-FDLA), and fluoro-2,4-dinitrophenyl-5-D-leucine amide (D-FDLA) were used with reference amino acids L-Ser, D-Ser, L-Hse, D-Hse, L-Glu, D-Glu, L-Thr, D-Thr, L-Val, D-Val, D-Pro, L-Pro, L-Tyr, D-Tyr, D-allo-Thr, L-Leu, D-Leu, L-Ile, D-Ile, D-allo-Ile, L-allo-Ile, D-Phe, L-Phe, (2S,4S)-4-hydroxy-Pro, (2R,4S)-4-hydroxy-Pro, (2S,4R)-4-hydroxy-Pro, L-Ser, D-Ser, (Sigma, Switzerland), N-methyl-L-Ile (hydrochloride, ABCR, Germany), L-allo-Thr (ICN Biomedicals, USA), (2R,4R)-4-hydroxy-Pro (Aldrich, USA), and (3S)- and (3R)-3-amino octanoic acid (ABCR, Germany).

Amino acid feeding experiment

Nostoc sp. UHCC 0702 and S. hofmannii PCC 7110 were cultivated in the presence of different Pro variants and hypothesized OHMePro synthesis intermediates to investigate if they are incorporated in the laxaphycin structures. Nostoc sp. UHCC 0702 was grown on Z8 media and S. hofmannii PCC 7110 was grown on Z8 media without added nitrogen. The media were modified by adding (2S,4S)-4-methyl-proline, (2R,4R)-4-methyl-proline, (2S,4R)-4-methyl-proline, (ABCR, Germany), (2S,4S)-4-hydroxyproline (Sigma, USA), (2S,4R)-4-hydroxyproline (Sigma, Japan), or L-Leu (Sigma, Switzerland) to concentrations of 10 μM. Racemic OHLeu was used at 40 μM concentration. Unaltered medium was used in control cultivation. Three replicates of each cultivation were grown and analyzed. The bacteria were grown in 100 mL Erlenmeyer flasks in 40 mL of medium for 17 days at 20–21 °C with photon irradiation of 15 μmol m $^{-2}$ s $^{-1}$. Cells were collected by centrifugation at 8000g for 5 min. Cells were frozen to −80 °C and freeze-dried. Freeze-dried biomass was weighed and extracted with 0.5 mL methanol and glass beads (0.5 mm glass beads, Scientific Industries Inc, USA) using a FastPrep cell disrupter two times for 25 s at a speed of 6.5 m s $^{-1}$. Samples were centrifuged at room temperature at 10 000g for 5 min and supernatant was collected for LC-MS analysis.

Stable isotope labeling

Nostoc sp. UHCC 0702 was labelled with ^{15}N in a nitrogen-free atmosphere in Z8 growth medium containing ^{15}N-urea as the sole source of nitrogen. The medium was bubbled with nitrogen-free argon with 20.9% O_2 and 0.45% CO_2 (quality 5.7; AGA Gas Ab, Sweden). The strain was grown for 21 days at 20 °C under photon irradiation of 15 μmol m $^{-2}$ s $^{-1}$. The biomass was collected and freeze dried. A total of 100 mg of the dried cells were extracted with methanol as described earlier and analyzed with LC-MS.

LC-MS

Freeze-dried bacterial cells were extracted with methanol as described earlier. Extracts and isolated heinamide fractions dissolved in methanol were analyzed with UPLC-QTOF (Acquity I-Class UPLC-SynaptG2-Si, Waters Corp., Milford, MA, USA) equipped with a Kinetex® C8 column (2.1 × 50 or 100 mm, 1.7 μm, 100 Å, Phenomenex, Torrance, CA, USA) injected with 0.5 or 1 μL of sample and eluted at 40 °C with 0.1% HCOOH in water (solvent A) and acetonitrile/isopropanol (1 : 1, +0.1% HCOOH, solvent B) at a flow rate of 0.3 mL min $^{-1}$. Different solvent gradients were used (Table S3†).

QTOF was calibrated using sodium formate and Ultramark® 1621, which gave a calibrated mass range from m/z 91 to 1921. Leucine Enkephalin was used at 10 s intervals as a lock mass reference compound. Mass spectral data were accumulated in positive electrospray ionization resolution mode. In MS/MS mode, Trap Collision Energy Ramp proceeded from 40.0 eV to 70.0 eV.

NMR spectroscopy

NMR spectra were obtained using a Bruker Avance III HD 800 MHz NMR spectrometer equipped with the TCI ^{1}H, ^{13}C, ^{15}N triple resonance cryoprobe. Data were collected at 30 °C in DMSO-d_6. In addition to ^{1}H and broadband-decoupled ^{13}C experiments, we employed two-dimensional TOCSY (TOtal Correlation SpectroscopY), DQF-COSY (Double Quantum Filtered COrrelation SpectroscopY), and EASY-ROESY (Efficient Adiabatic SYmmetrized Rotating-frame Overhauser Effect SpectroscopY)[21] experiments, and ^{13}C HSQC, ^{15}N HSQC, edited ^{13}C HSQC (heteronuclear Single Quantum Coherence), and ^{13}C HMBC (Heteronuclear Multiple Bond Correlation) experiments for the assignment of ^{1}H, ^{13}C, and ^{15}N chemical shifts. The 2D TOCSY was acquired with isotropic mixing time of 60 ms (10 kHz RF field). The mixing time for 2D ROESY experiment was 200 ms (RF field strength 4 kHz). For observing long-range H–C connectivities, the ^{13}C HMBC experiment was measured using $^{n}J_{CH}$ transfer time optimized for 8 Hz couplings. NMR experiment parameters are presented in Table S4.†

Genome sequencing and gene cluster analysis

Cultures were grown in 250 mL of modified Z8 medium lacking a source of combined nitrogen at 20–21 °C with photon irradiation of 15 μmol m $^{-2}$ s $^{-1}$. Half-liter Erlenmeyer flasks were used with 250 mL of medium, with constant sterilized air bubbling for 3 weeks. The fresh Nostoc sp. UHC 0702 cells were disrupted by bead beating in GOS buffer and DNA was extracted with the phenol–chloroform method as previously described.[20] Sequencing of the extracted DNA was per-

formed with PacBio RSII and MGI DBNSeq-G400 sequencing and assembled with HGAP3 (smrtportal 2.3.0).

The Nostoc sp. UHCC 0702 complete genome data was analyzed with AntiSMASH 5.0[22] and AntiSMASH 4.1[23] to identify the lxa biosynthetic gene cluster. BLASTp and CDD database searches were used to assign a predicted function to the proteins encoded in the lxa biosynthetic gene clusters from the Nostoc sp. UHCC 0702 and S. hofmannii PCC 7110 genomes. The condensation domain of $LxaC1_3$ was analyzed with Natural Product Domain Seeker (NaPDoS)[24] and compared with other known Dhb-related condensation domains $LxaC1_3$, $HasO_2$, $NdaA_1$, and $PuwF_2$, which are involved in the biosynthesis of scytocyclamides,[10] hassalladins,[25] nodularins,[20] and puwainaphycins,[26] respectively.

Heterologous expression of lxaN gene

The function of LxaN, a putative proline 4-hydroxylase, was tested through heterologous expression in Escherichia coli BL21(DE3). LxaN was expected to hydroxylate available Pro in the E. coli cells. Plasmid constructs with the lxaN gene were prepared and transformed to E. coli, and amino acid analysis was performed to detect 4-OHPro in the transformed cells.

Genomic DNA was extracted from Nostoc sp. UHCC 0702 as previously described. The lxaN gene was amplified by PCR with primers LxaN-F (5 -gtggtggtgctcgagtgcggccgcaTTAAATAAGAACTTTGTCCAATAG-3) and LxaN-R (5 -ggacagcaaatgggtcgcggatccgATGTCCTATACCAATCAAAC-3). The vector pET28a (+) was digested with restriction enzymes EcoRI and HindIII (Promega, USA). The lxaN gene was inserted into pET28a (+) using the NEBuilder® HiFi DNA Assembly Cloning Kit (New England BioLabs, USA) to create plasmid pET28a-LxaN. To allow inducible expression, the lxaN gene was placed behind a T7 promoter and lac operator. The plasmid pET28a-LxaN was transformed into E. coli BL21 (DE3). A negative control was prepared by transforming an empty pET28a plasmid into E. coli BL21 (DE3). Clones were selected by using LB agar plates containing 50 g mL ^{1} kanamycin, 1 g 100 mL ^{1} arabinose, and 1 g 100 mL ^{1} glucose. Three transformants and a negative control were transferred to liquid LB medium with 50 g mL ^{1} kanamycin and 1 g 100 mL ^{1} glucose and incubated at 37 °C overnight with shaking (170 rpm). Aliquots of these primary cultures were used to inoculate 10 mL of fresh LB medium supplemented with selective antibiotics, 1 g 100 mL ^{1} arabinose, and 1 g 100 mL ^{1} glucose. The cultures were incubated at 37 °C with shaking at 170 rpm until OD_{600} of 0.6 was achieved. The cultures were induced with 0.1 mM isopropyl -D-1-thiogalactopyranoside (IPTG) and incubated at 18 °C for 48 h with shaking (200 rpm).

Biomass was collected from the growth culture by centrifugation and the supernatant was discarded. The biomass was frozen at 80 °C in screw-cap tubes. To extract biomass, 200 L of 0.5 mm glass beads (Scientific Industries Inc, USA) and 1 mL 100% methanol were added into the tube. The mixture was homogenized by using Fastprep-24 twice for 20 seconds at a speed of 6 m s ^{1}. Cell debris was pelleted by centrifugation at 13 400g for 5 min. Supernatant was transferred into new tubes and dried with nitrogen gas flow. Amino acid analysis was performed with Marfey's method as previously described with UPLC-QTOF mass spectrometry,[20] using reference amino acids (2S,4S)-4-hydroxyproline, (2S,4R)-4-hydroxyproline, (2R,4S)-4-hydroxyproline (Sigma, Switzerland), and (2R,4R)-4-hydroxyproline (Aldrich, USA).

Results

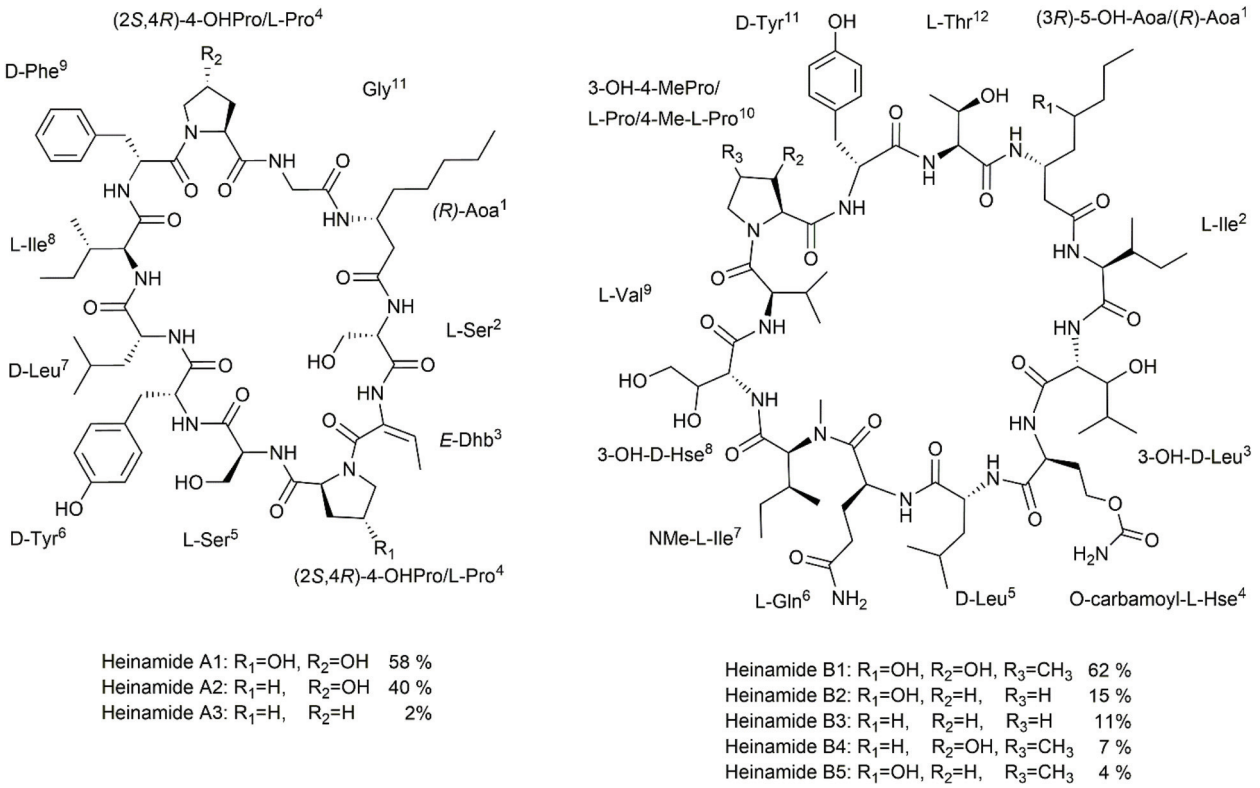
Antimicrobial screening

A methanol extract of Nostoc sp. UHCC 0702 inhibited the growth of A. flavus FBCC 2467. The other tested fungal and bacterial strains were not a ected by the extract.

Structure of heinamides

Eight peptides belonging to the laxaphycin family were identified from Nostoc sp. UHCC 0702 and named heinamides A1–A3 and B1–B5 (Fig. 1). Heinamides A1–A3 are 11-residue laxaphycins and heinamides B1–B5 are 12-residue laxaphycins. The 11-residue heinamides follow the amino acid sequence Aoa^{1}-Ser^{2}-Dhb^{3}-((2S,4R)-4-OHPro/Pro)^{4}-Ser^{5}-Tyr^{6}-Leu^{7}-Ile^{8}-Phe^{9}-((2S,4R)-4-OHPro/Pro)^{10}-Gly^{11}; the di erence between variants was the hydroxylation of Pro^{4,10}. The 12-residue heinamides have the amino acid sequence (Aoa/5-OH-Aoa)^{1}-Ile^{2}-OHLeu^{3}-(O-carbamoyl-Hse)^{4}-Leu^{5}-Gln^{6}-(N-Me-Ile)^{7}-OHHse^{8}-Val^{9}-(OHMePro/Pro/4-MePro)^{10}-Tyr^{11}-Thr^{12}, where the di erences between variants is in the hydroxylation of the octanoic acid^{1} and hydroxylation and methylation of the Pro^{10}. The stereochemistry of the compounds is derived from stereospecific amino acid analysis (Table S11†) and NMR.

The methanol extract from Nostoc sp. UHCC 0702 was analyzed with UPLC-QTOF to determine the initial heinamide structures. Cultivation on ^{15}N-containing medium and comparison of the mass data of ^{15}N-labeled and unlabeled compounds showed the presence of 11 nitrogen atoms in 11-residue heinamides A1–A3 and 14 nitrogen atoms in 12-residue heinamides B1–B5 (Table S5, Fig. S2†). The product ion spectra of protonated heinamides consisted of many evenly intense ions, which is typical for some cyclic peptides (Fig. S3–S6†). The best continuous data for amino acid sequencing started as expected from the amino acids next to the Pro N-terminus (red markings in Fig. S3 and S4†). The structure of heinamides A3, B4 and B5 is proposed solely on the basis of analyses of their LC MS/MS data. The structures assigned by LC-MS were consistent with those assigned using NMR. ^{1}H, ^{13}C, DQF-COSY, TOCSY, EASY-ROESY, ^{13}C HSQC, edited ^{13}C HSQC, ^{15}N HSQC, and ^{13}C HMBC NMR spectra were obtained from the purified heinamides A2 and B2 and heinamide mixtures A1 : B3 : B4 (7 : 2 : 1) and B1 : B2 (1 : 1). NMR spectra are presented in Fig. S7–S10† and numerical data with COSY, ROESY, and HMBC correlations in Tables S6–S10.† COSY, HMBC, and ROESY uninterrupted correlation chain specified the subunit sequences in both 11-residue (HA A1) and 12-residue (HA B2) heinamides (Fig. S11 and S12†). Heinamide structures were in good agreement with previously described laxaphycin common structural features (Table S1†).



Heinamide A1: R_1=OH, R_2=OH 58 %
Heinamide A2: R_1=H, R_2=OH 40 %
Heinamide A3: R_1=H, R_2=H 2%

Heinamide B1: R_1=OH, R_2=OH, R_3=CH_3 62 %
Heinamide B2: R_1=OH, R_2=H, R_3=H 15 %
Heinamide B3: R_1=H, R_2=H, R_3=H 11%
Heinamide B4: R_1=H, R_2=OH, R_3=CH_3 7 %
Heinamide B5: R_1=OH, R_2=H, R_3=CH_3 4 %

Fig. 1 Chemical structures of heinamides and relative intensities of their MS signals (%), puri ed from Nostoc sp. UHCC 0702.

In heinamides A1–A3, the only new element was (2S,4R)-4-OHPro or L-Pro in position 10. Two di erent ppm value sets for 4-OHPro were present for HA A1. In amino acid analysis (2S,4R)-4-OHPro was the only 4-OHPro enantiomer present of the four possible alternatives (Tables S6 and S11†). More new elements were found in 12-residue heinamides B1–B5. The hydroxy group in the -amino octanoic acid[1] has not been described earlier in laxaphycins, but COSY, HSQC, and HMBC correlations show the presence of a methine ($_H$ = 3.46, $_C$ = 66.2 ppm) group in position 5, which is most probably bonded to oxygen in HA B1 (Fig. S9F, I and N and S12†). In many 12-residue laxaphycins, Leu[5] is hydroxylated but not in heinamides according to NMR (Fig. S9E, L and M†). In previously described 12-residue laxaphycins, the amino acid in position 8 has almost exclusively been 3-OHAsn. However, the amino acid is 3-hydroxy-homoserine in heinamides B1–B5 (Tables S8–S10, Fig. S9 and S10†). Lastly, NMR data showed another Hse is in position 4 in 12-residue heinamides. Furthermore, COSY, HSQC, and HMBC correlations showed that this subunit was actually O-carbamoylated (Fig. S9, especially frame O, Tables S8–S10†). Isobaric 3- or 4-OHGln were ruled out as the C3 and C4 were methylenes according the edited ^{13}C-HSQC (Fig. S9,† frame I). Product ion spectra of protonated heinamides B1–B5 show the loss of the carbamoyl group as carbamic acid (61.02 Da) until the loss of the O-carbamoyl-Hse[4] subunit itself (Fig. S4–S6†). No 12-residue heinamide without an O-carbamoyl group in Hse was found.

Antifungal activity

After fractionation of the extract and identification of the products, disc di usion assays were performed with the purified compounds. Inhibition of fungal growth was observed with peptides of a single laxaphycin type as a hazy inhibition zone. Synergy was observed between 11- and 12-residue compounds as a clear inhibition zone (Fig. 2).

Predicted heinamide biosynthetic pathway from Nostoc sp. UHCC 0702

We sequenced the complete genome of Nostoc sp. UHCC 0702 (GenBank accession number CP071065). The genome is a

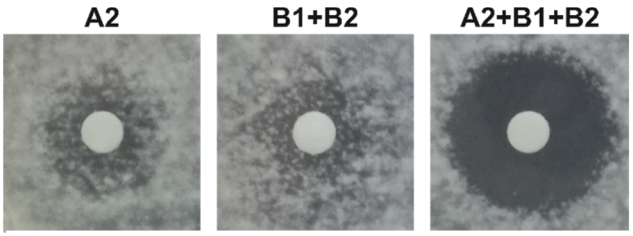


Fig. 2 Inhibition of growth of Aspergillus avus by heinamides. Heinamide A2 (200 µg), heinamide B1 + B2 (200 µg), and heinamide A2 with B1 + B2 (100 µg + 100 µg). A mix of B1 and B2 was used due to imperfect puri cation of the compounds, with the fraction including both B1 and B2 (1 : 1). Individual compounds exhibit weak inhibition and synergistic activity is seen when combining 11- and 12-residue molecules.

single 8587 404 bp circular chromosome without extrachromosomal elements. Sequence coverage with PacBio RSII was 90x and with MGI DBNSeq-G400 46x. Bioinformatics analysis with AntiSMASH identified 13 putative NRPS/PKS pathways. Comparison with the S. hofmannii PCC 7110 lxa biosynthetic gene cluster and chemical structures of the identified heinamides confirmed the presence of a lxa biosynthetic gene cluster in Nostoc sp. UHCC 0702 (Fig. 3).

The cluster is 93 kb long and encodes 13 open reading frames annotated as ORF1–2, lxaI1, lxaJ1, lxaK1, lxaA2, ORF3, lxaB, lxaC1, lxaE, lxaF, lxaG, and lxaM (Fig. 3, Table 1). lxaC1 is an NRPS with adenylation domain binding pockets matching the amino acid sequence of 11-residue heinamides (Table S12†). lxaC1 is flanked with PKSs lxaB and lxaE followed by cupin-like domains lxaFGM (Fig. 3, Table 1). The Nostoc sp. UHCC 0702 genes lxaI1, lxaJ1, and lxaK1 adenylation domain binding pockets match the amino acid sequence of 12-residue heinamides (Fig. 3, Table S12†). The position of epimerase domains in the lxa biosynthetic enzymes match the positions of D-amino acids in the elucidated structures (Fig. 1 and 3). Laxaphycin biosynthesis is predicted to be initiated by an activating enzyme lxaA with FAAL and ACP domains (Fig. 3). However, such an enzyme was not encoded in the Nostoc sp. UHCC 0702 lxa biosynthetic gene cluster. Instead, an ACP in the cluster with 41% sequence identity in BLASTp to the ACP domain of LxaA was identified and designated LxaA2 (Fig. 3, Table 1).

The lxa biosynthetic gene cluster was missing the essential FAAL domain for the initiation of the pathway and the genes to synthesize the modified amino acids (2S,4R)-4-OHPro, OHMePro, O-carbamoyl homoserine, and an ABC-transporter. The LxaA FAAL domain of S. hofmannii PCC 7110 was used as BLASTp query to search for the initiating domain, and a FAAL with the highest identity was annotated LxaA1 (Fig. 3, Table 1). This gene was located 1.6 Mb upstream from the lxa biosynthetic gene cluster (Fig. 3). In S. hofmannii PCC 7110, the LxaA protein includes two domains, a FAAL and an ACP domain (Fig. S16†). Together LxaA1 and LxaA2 act as LxaA (Fig. 3).

We predict that the cupin-like domain proteins LxaF, LxaG, and LxaM hydroxylate Leu[3], Hse[8], and Aoa (Fig. 3). ORF1–3 were found in the gene cluster with no predicted function. Eleven-residue heinamides contain the non-proteinogenic amino acid Dhb[3], the dehydration product of Thr. We discovered that the condensation domain $LxaC1_3$ groups in the modified AA clade of condensation domains that act in the dehydration of Thr and Ser from the previous module to Dhb or Dha (Fig. S13†). Three putative carbamoyltransferases were found in the genome, but none were assigned to the gene cluster. With the lack of known carbamoyltrasferases acting on amino acids, no reliable prediction could be made for the enzyme responsible for the O-cabamoylation of HSe.

A set of genes encoding enzymes homologous to genes producing (2S,4S)-4-methylproline in cyanobacterial metabolites were identified 389 kb downstream from the lxa biosynthetic gene cluster (Fig. 3, Table 1). These enzymes are a L-Leu 5-hydroxylase (LxaO), a zinc-binding dehydrogenase (LxaP), and a pyrroline-5-carboxylate reductase (LxaQ) (Fig. 3, Table 1). Flanking lxaOPQ were two genes lxaN and lxaR that encode -ketoglutarate-dependent oxygenases (Fig. 3, Table 1). LxaN was discovered to have a homolog also in the S. hofmannii PCC 7110 genome 11 kb downstream of the scytocyclamide gene cluster (WP_017742662.1). LxaN enzymes from Nostoc sp. UHCC 0702 and S. hofmannii PCC 7110 share 93% amino acid sequence identity. LxaN belongs to the pfam05721 class of oxygenases. We predicted that LxaN hydroxylates L-Pro to (2S,4R)-4-OHPro found in heinamide and scytocyclamide structures and that LxaR acts in OHMePro production hydroxylating the 3-carbon (Scheme 1). LxaR belongs to the pfam13640 class of -ketoglutarate-FeII dependent oxygenases known to hydroxylate amino acids.

LxaN of Nostoc sp. UHCC 0702 was heterologously expressed in E. coli to assess if it hydroxylates L-Pro to (2S,4R)-4-OHPro. Amino acid analysis performed with Marfey's method showed that a derivatized extract from E. coli with the LxaN construct matched the retention time of derivatized (2S,4R)-4-OHPro (Fig. 4). A control E. coli strain without added

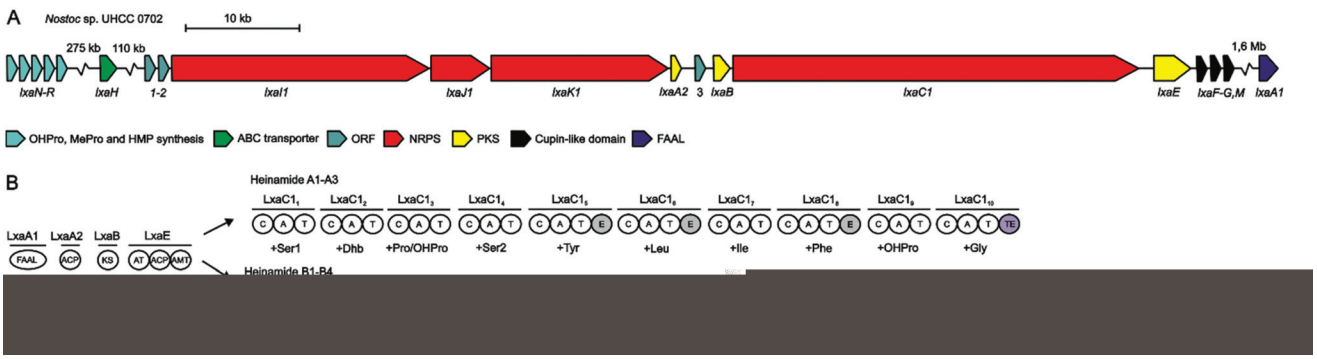


Fig. 3 The laxaphycin (lxa) biosynthetic gene clusters and putative biosynthetic scheme in Nostoc sp. UHCC 0702. (A) Organization of predicted heinamide biosynthetic genes. (B) Proposed biosynthetic pathway of heinamides. NRPS non-ribosomal peptide synthetase, PKS polyketide synthase, FAAL fatty acyl AMP Ligase, ACP acyl carrier protein, KS ketosynthase, AT acyltransferase, AMT aminotransferase, C condensation domain, A adenylation domain, T thiolation domain, M methylation domain, TE thioesterase domain.

